The Use of Machine Learning Algorithms and Statistical Models to Classify Aphasia Severity

BY

Tea Kristiane Espeland Uggen

Thesis submitted in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Under the supervisor of

A./Prof. Tapan Rai, A./Prof. Erin Godecke and Prof. Louise Ryan

UNIVERSITY OF TECHNOLOGY SYDNEY

SCHOOL OF MATHEMATICAL AND PHYSICAL SCIENCES

SEPTEMBER 2020

Certificate of Authorship

I, Tea Kristiane Espeland Uggen, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Science, at the University of Technology Sydney (UTS).

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research is supported by an Australian Government Research Training program.

Production Note:

Signature: Signature removed prior to publication.

Date: 03/09 /2020

Acknowledgements

I wish to extend my deepest appreciation for my primary supervisor, A./Prof. Tapan Rai, for the opportunity to work on the most interesting and rewarding research project, for all your guidance and for the substantial amount of feedback you have given me over the past four years. I can see a distinctive improvement in my way of thinking, writing and presenting as a researcher from four years ago and this is largely thanks to you and the faith you have placed in me. Thank you for taking me on as your student and for always pushing me forward. I would also like to express my gratitude to my co-supervisors, A./Prof. Erin Godecke and Prof. Louise Ryan, for supporting me in my research, sharing your knowledge and providing me with numerous opportunities to share my research, collaborate and grow as a researcher.

A Ph.D. is no easy job, especially not without a supportive network of friends and colleagues. Thankfully I was lucky enough to hit the goldmine of friends in my fellow Ph.D. and postdoc colleagues. My most heartfelt thanks go out to our UTS statistics group for your friendship, your help with R and LaTex bugs, your assistance in presentation preparations and last but most importantly your emotional support during the downs of the emotional rollercoaster that is a Ph.D. A special shout out to Marianne, Hon, Luca, James, Wilson, Gioia, Bojana and Mahrita for always being there for me.

I want to thank the mathematics and statistics staff at UTS as well for taking the time to provide me with feedback throughout my degrees at UTS and for being such a knowledgeable and approaching team. I am very fortunate to have been surrounded by such great teachers and mentors. An especially big thanks to Dr. Stephen Bush and Prof. Matt Wand. Thank you as well to the helpful people at UTS Hacky Hour for helping me understand and work through my coding bugs the numerous times I've been stuck on a problem - I would not have made it through without your help.

I am also deeply grateful for my Australian family, uncle Erik, aunt Barbara and Ross, for your valuable input, encouragement and support during my studies.

Finally, I want to express my deepest thanks to my family in Norway, who have supported me during my nine years of studying in Sydney. In particular - my wonderful mom and dad for their constant support during my studies, for following my progress for every subject completed and for assisting as my non-technical audience for countless essays, presentations and articles. Furthermore, to the most encouraging and caring brother and sister in the world - thanks for listening to my Ph.D. rants and for all your words of encouragement. I also want to thank my two grandmothers for being the best, toughest and most inspiring role-models. Only the drive and courage passed on to me from you both could have gotten me through these past nine years. Thanks for all the knowledge and life experiences you have shared with me.

Abstract

Aphasia is a communication impairment usually caused by stroke, identified in approximately one third of stroke survivors. Determining the severity level for patients with aphasia is a necessary pre-cursor in determining the optimal rehabilitation pathway for each patient. Language assessments are used to evaluate aphasia severity. There are numerous existing language assessments in the aphasia literature. In this thesis, we will focus on two common assessments: the Western Aphasia Battery-Aphasia Quotient (WAB-AQ) and Discourse Analysis (DA). The WAB-AQ is considered a gold standard of aphasia assessments. However, it is lengthy and strenuous to undertake. DA is a shorter and more manageable assessment. However, there are no existing standards to classify aphasia severity from this assessment. Our research aims to determine whether it is possible to use the DA assessment to predict the aphasia severity level of the WAB-AQ. There are three research questions in this thesis: 1) Is it possible to use machine learning (ML) and natural language processing (NLP) algorithms to automatically identify the levels of aphasia severity from patients' speech transcripts, recorded during DA assessments? This objective is achieved by developing an NLP algorithm to convert the speech transcripts to numerical measures of speech; then performing classification using these measures to predict WAB-AQ severity level with various ML techniques. We extend the binary confusion matrix to a multi-class confusion matrix and develop associated model assessment metrics, specifically developed to tackle our three-class classification problem, to evaluate the predictive power of our models. We also determine the measures of speech from the NLP algorithm which are most important in classifying aphasia severity by developing an accuracy-based feature-selection algorithm able to incorporate multiple machine learning techniques. 2) How does the performance of our proposed method compare to a comparative baseline method? This baseline method uses existing clinical measures of DA as predictors in classification models to predict WAB-AQ severity level. 3) How do existing measures of DA compare among each other in their ability to predict aphasia severity? We use the measures individually as model predictors and compare the model assessment metrics from each model. Our results show that our proposed approach can be used to identify aphasia severity in post-stroke patients instantly with up to 73.6% accuracy. Furthermore, our findings indicate that our proposed method is superior in predicting aphasia severity level than the baseline method, with the NLP models yielding consistently higher model accuracies than the baseline models.

Keywords: *aphasia, classification, machine learning, natural language processing, severity*

Contents

List of Figures

List of Tables